

Search Notes

132113

Schreiber, David

From: Sitton, Jehanne Souaya
Sent: Monday, August 30, 2004 12:39 PM
To: Schreiber, David
Subject: score/length sequence search request for 10/007781

please perform the following search for application 10/007,781: in nucleic acid databases except pending database.

please perform a score/length sequence search for complements of SEQ ID NO: 1, residues 3930-3970, with a minimum hit length of 11 and a maximum hit length of 50, and a score over length cutoff at 70%.

please provide all hits, or at least the first 200.

I've never asked for this type of search before, so I have a question:
The search I've asked for is over a small portion of the nucleic acid. If the length I've asked for decreases the time for the search and the processing time for the hits, and if the search I've asked for doesn't cover the complements of the antisense sequence, could I please also have a search of the complements of the antisense sequence of SEQ ID NO: 1, with the same parameters as above. if doing this extra search will take too much time, I can make do without it.

thanks,
Jehanne Sitton
Primary Examiner
Art Unit 1634
Rem-2D81 (office)
Rem-2C70 (mailbox)
703-272-0752

SCORE OVER LENGTH SEARCHES

Attached is a score over length search. This search was developed to overcome limitations in most standard search systems which favor large sequences with high scoring, but lesser overall identity over smaller sequences with higher overall identity. This search is especially useful for relatively small nucleic acid or polypeptide target sequences (antisense, fragments, probes, primers, RNAi, epitopes, haptens, etc.) claimed functionally via a form of hybridization and/or identity language and having defined upper and lower polynucleotide and or polypeptide length limits.

The score over length search is performed by first running the query sequence using examiner-specified identity and polynucleotide or protein length limit parameters, and saving 65,000 hits and 0 alignments from each desired database. The resulting output is reformatted using a Microsoft Word macro and is imported into Excel. The summary table data are then sorted by the ratio of score of each hit sequence divided by its length and the accession numbers for all hits below the examiner's desired score over length parameters are deleted. The remaining accession numbers are used to pull the corresponding sequences from the databases into subdatabases enriched for good hits and the query sequence is re-run against these subdatabases to yield the final results.

The score over length cutoff for this search is 70%

Examiner Please Note: This cover sheet should be included when submitting results to be scanned.